INTEGRATED STOCHASTIC INVESTMENT & OPERATIONS STRATEGY

Fernanda Thomé[1] and Camila Metello[1]

[1]PSR Inc.

June 2016

Abstract

Many capacity planning models used today are based on a Benders decomposition scheme [1, 2] composed of: (i) a MIP-based "investment module" which determines a trial expansion plan; (ii) a SDDP-based "operation module" which calculates the expected operation costs for the trial plan; and (iii) Benders cuts from the operation to the investment module, whose coefficients are calculated from the expected marginal costs of the capacity constraints in the operation module at the optimal solution.

Although this "traditional" planning model has been successfully applied in many countries, it has an inherent limitation, which has become more significant with the penetration of renewables with short construction times, such as solar: the optimal expansion plan is "static", i.e. investment decisions do not change as the system state evolves (e.g. load growth is lower than expected, a very rainy season occurs etc.). As a consequence, there is a growing interest in the calculation of an integrated stochastic investment & operations strategy.

This paper describes an extension of the SDDP algorithm [3] that allows the calculation of this integrated strategy. The first (and obvious) step of this extension is to represent investment decisions as state variables in the SDDP recursion. The second step is to represent the construction time of each candidate project in the recursion; this requires an efficient modeling of time delays in the update of state variables. The final step is to represent the integrality of investment decisions in the multistage stochastic optimization scheme. This is done by applying a customized Lagrangian scheme to the scheduling/investment subproblem of each stage and scenario that produces the strongest possible convex cut to the previous stage's future cost function.

The application of the proposed algorithm will be illustrated in realistic capacity planning studies of the Central America system.

1 "Traditional" G&T Expansion Planning Model

Figure (1) illustrates the main components of a generation/transmission capacity planning model which represents the system operation by multistage stochastic programming techniques (in this case, the SDDP algorithm).

1.1 Investment module

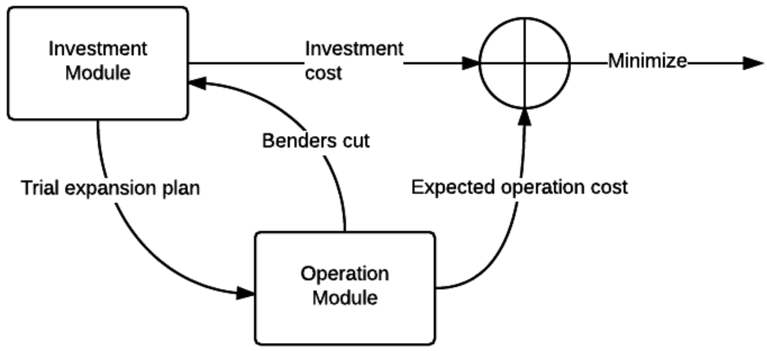
The binary variables $x_{t,i}$, $x_{t,j}$, $x_{t,r}$ and $x_{t,k}$ represent respectively the construction of candidate projects (hydro i, thermal j, renewable r and circuit/transformer k) in stage t. The sets of candidate projects are represented as $\mathbb{I}_x$, $\mathbb{J}_x$, $\mathbb{R}_x$ and $\mathbb{K}_x$. In turn, the sets of existing generation/transmission devices are $\mathbb{I}_e$, $\mathbb{J}_e$, $\mathbb{R}_e$ and $\mathbb{K}_e$.

The investment→operation Benders decomposition iterations are indexed by $m = 1, \ldots, M$. We show below the investment problem after M iterations.

$$\text{Min} \sum_t \sum_{i \in \mathbb{I}_x} I_i x_{t,i} + \sum_{j \in \mathbb{J}_x} I_j x_{t,j} + \sum_{r \in \mathbb{R}_x} I_r x_{t,r} + \sum_{k \in \mathbb{K}_x} I_k x_{t,k} + w \tag{1a}$$

$$x_{t,i} \geq x_{t-1,i} \qquad t = 2, \ldots, T; i \in \mathbb{I}_x \tag{1b}$$

Figure 1: Components of a "traditional" capacity expansion planning model



$$x_{t,j} \geq x_{t-1,j} \qquad t = 2, \ldots,\ T; j \in \mathbb{J}_x \tag{1c}$$

$$x_{t,r} \geq x_{t-1,r} \qquad t = 2, \ldots,\ T; r \in \mathbb{R}_x \tag{1d}$$

$$x_{t,k} \geq x_{t-1,k} \qquad \forall t = 2, \ldots,\ T; k \in \mathbb{K}_x \tag{1e}$$

$$w \geq \sum_t \left(\sum_{i \in \mathbb{I}_x} \overline{\mu}^m_{t,i} x_{t,i} + \sum_{j \in \mathbb{J}_x} \overline{\mu}^m_{t,j} x_{t,j} + \sum_{r \in \mathbb{R}_x} \overline{\mu}^m_{t,r} x_{t,r} + \sum_{k \in \mathbb{K}_x} \overline{\mu}^m_{t,k} x_{t,k} + \ \overline{\mu}^m_{0t} \right) \quad \mathrm{m} = 1, \ldots,\ M \tag{1f}$$

1.2 Operation module

Given the trial optimal investment decisions $\left\{x^*_{t,i}\right\}$, $\left\{x^*_{t,j}\right\}$, $\left\{x^*_{t,r}\right\}$ and $\left\{x^*_{t,k}\right\}$ of the investment module in the M-th iteration of the Benders decomposition scheme, we solve the stochastic scheduling problem using the SDDP algorithm, described next.

1.3 SDDP Formulation

1.3.1 Notation

Indices

$t\ =\ 1,\ \ldots,\ T$	time stages (typically weeks or months)
$\tau\ =\ 1,\ \ldots,$	intra-stage time blocks (e.g. peak/medium/low demand or 168 hours in a week)
$s\ =\ 1,\ \ldots, S$	scenarios for each stage t produced by the stochastic models (typically inflows and renewable generation; also loads, equipment availability and fuel costs)
$l\ =\ 1,\ \ldots, L$	set of scenarios for stage $t+1$ conditioned to scenario s in stage t
$i\ =\ 1,\ \ldots, I$	storage devices (typically hydro plants; also fuel storage, batteries, emission limits and some types of contracts)
$m \in M_i$	set of hydro plants immediately upstream of plant i
$j\ =\ 1,\ \ldots, J$	dispatchable devices (typically, thermal plants; also, some controllable renewables and price-responsive demand)
$r\ =\ 1,\ \ldots, R$	non-dispatchable devices (typically, wind, solar and biomass)
$n\ =\ 1,\ \ldots, N$	transmission network buses
$k\ =\ 1,\ \ldots, K$	transmission network components (circuits, transformers and FACTS devices such as phase shifters and smart wires)

$p = 1, \ldots, P$	number of hyperplanes (Benders cuts) in the future cost function

Decision variables for the operation problem in stage t, scenario s

$v_{t+1,i}$	stored volume of hydro i by the end of stage t
$u_{t,i}$	turbined volume of hydro i stage t
$\nu_{t,i}$	spilled volume of hydro i in stage t
$e_{t,\tau,i}$	energy produced by hydro i in block τ, stage t
$g_{t,\tau,j}$	energy produced by thermal plant j in block τ, stage t
α^l_{t+1}	present value of expected future cost from $t+1$ to T conditioned to scenario l in $t+1$

Known values for the operation problem in stage t, scenario s

$\hat{a}^s_{t,i}$	lateral inflow to hydro i in stage t, scenario s ($\hat{a}^s_t$ set of inflows for all hydro plants)
$\hat{v}^s_{t,i}$	stored volume of hydro i in the beginning of stage t, scenario s ($\hat{v}^s_t$ set of stored volumes for all hydro plants)
$\overline{v}_i$	maximum storage of hydro i
$\overline{u}_i$	maximum turbined outflow of hydro i
ϕ_i	production coefficient (kWh/m^3) of hydro i
$\overline{g}_j$	maximum generation of thermal plant j
c_j	variable operating cost of thermal plant j
$\hat{r}^s_{t,\tau,r}$	energy produced by renewable plant r in stage t, block τ, scenario s
$\hat{d}_{t,\tau}$	demand of block τ, stage t

Multipliers

$\pi^h_{t,i}$	multiplier of the storage balance equation of hydro i (see problem formulation)
$\pi^a_{t,i}$	multiplier of the conditioned inflow equation of hydro i (see problem formulation)

p^{th} **Benders cut coefficients**

$\hat{\varphi}^{hp}_{t+1,i}$	coefficient of hydro plant i's storage, $v_{t+1,i}$
$\hat{\varphi}^{\text{ap}}_{t+1,i}$	coefficient of hydro plant i's inflow, $a^l_{t+1,i}$
$\hat{\varphi}^{op}_{t+1}$	constant term

Stochastic streamflow model coefficients

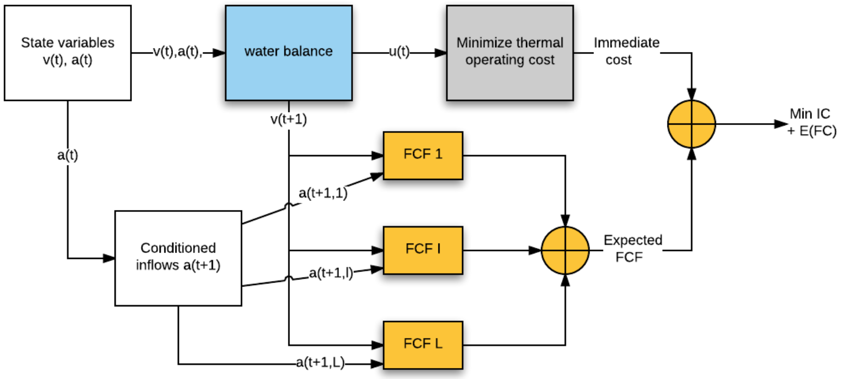
$\hat{\mu}_{t,i}$	mean of the lateral inflow to hydro i in stage t.
$\hat{\sigma}_{t,i}$	standard deviation of the lateral inflow to hydro i in stage t.
$\hat{\rho}_{t,i}$	serial correlation of the lateral inflow to hydro i in stage t.

1.4 Problem Formulation

Figure (2) shows the main components of the operation problem for stage t, scenario s:

1. SDDP *state variables* at the beginning of the stage (in this example, initial storage $v(t)$ and inflow along the stage, $a(t)$);
2. *reservoir storage balance* equations, which determine the hydro turbined outflow, $u(t)$;

Figure 2: Main components of SDDP's operation problem for stage t, scenario s



3. *power balance* equation, which determines the least-cost operation of the thermal plants required to meet the residual load (after subtracting hydro generation and renewable production). In the SDDP formulation, the resulting operation cost is known as the *immediate cost function* (ICF);

4. *future cost functions* (FCF) $l = 1, \ldots, L$ of the SDDP state variables for the next stage: the final storage $v(t+1)$ and $l = 1, \ldots, L$ conditioned inflow scenarios $a(t+1, l)$.

5. the objective function: minimize the sum of immediate cost ICF and the mean future cost $\frac{1}{L} \sum \text{FCF}_l$

Objective function

$$\alpha_t \left(\hat{v}_t^s, \hat{a}_t^s\right) = Min \sum_j c_j \sum_\tau g_{t,\tau,j} + \frac{1}{L} \sum_l \alpha_{t+1}^l \tag{2a}$$

Storage balance for each stage

$$v_{t+1,i} = \hat{v}_{t,i}^s + \hat{a}_{t,i}^s - \left(u_{t,i} + \nu_{t,i}\right) + \sum_{\sqcap \in \mathcal{U}_i} \left(u_{t,\sqcap} + \nu_{t,\sqcap}\right) \qquad \leftarrow \pi_{t,i}^h \tag{2b}$$

Note: for notational simplicity, we do not represent in this formulation real-life features of the storage balance equations such as evaporation, filtration, water diversion for irrigation and city supply, transposition and others.

Storage limits

For the existing hydro plants, the storage limit is a given value, $\overline{v}_i$

$$v_{t+1,i} \leq \overline{v}_i \qquad \forall i \in \mathbb{I}_e \tag{2c}$$

For the candidate hydro plants, the storage limit $\overline{v}_i^*$ depends on the investment decision $x_{t,i}^*$ at the current iteration of the Benders decomposition scheme (investment module):

$$v_{t+1,i} \leq \overline{v}_i^* \left(= \overline{v}_i \times x_{t,i}^*\right) \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t,i}^{\overline{v}} \tag{2d}$$

Note that the SDDP operating module "sees" both existing and candidate storage limits $\overline{v}_i$ and $\overline{v}_i^*$ as given values. In other words, the operating module does not "know" that is being run as part of a Benders decomposition scheme with the investment module. The investment information is only used explicitly in the calculation of the Benders cuts from the operation to the investment module. One advantage of this scheme that the same model used in operations studies can be used in planning studies, without any modification.

Turbined outflow limits

$$u_{t,i} \leq \overline{u}_i \qquad \forall i \in \mathbb{I}_e \tag{2e}$$

$$u_{t,i} \le \overline{u}_i^* \left(= \overline{u}_i \times x_{t,i}^*\right) \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t,i}^{\overline{u}} \tag{2f}$$

The same logic of the storage limit (2d) applies to the turbined outflow (2f).

Hydro generation

$$e_{t,i} = \phi_i u_{t,i} \tag{2g}$$

For notational simplicity, hydro generation is represented here as a linear function of the turbined outflow. In real-life applications, other factors are modeled such as the variation of reservoir head with storage, increase of tailwater with total outflow, encroachment from the downstream reservoirs.

$$\sum_\tau e_{t,\tau,i} = e_{t,i} \tag{2h}$$

$$e_{t,\tau,i} \le \overline{e}_i \tag{2i}$$

Note that it is not necessary to have investment variables for hydro energy production because this is already done for turbined outflow.

Thermal generation

As in the hydro case, the generation capacity of candidate plants changes with the investment module iterations.

$$g_{t,\tau,j} \le \overline{g}_j \qquad \forall j \in \mathbb{J}_e \tag{2j}$$

$$g_{t,\tau,j} \le \overline{g}_j^* \left(= \overline{g}_j \times x_{t,j}^*\right) \quad \forall j \in \mathbb{J}_x \quad \leftarrow \pi_{t,\tau,j}^{\overline{g}} \tag{2k}$$

As in the previous cases, we show a simple representation of thermal generation in this formulation. In actual applications, there are models for efficiency curves, multiple fuels, fuel storage and unit commitment.

Renewable generation

Renewable generation is represented as energy production scenarios $\left\{\hat{r}_{t,\tau,r}^s\right\}$ in the power balance equations, described next.

Transmission network equations and constraints

The first set of equations represents the power balance in each bus (Kirchhoff's first law)

$$Sf_{t,\tau} + e_{t,\tau} + g_{t,\tau} = \hat{d}_{t,\tau} - r_{t,\tau}^{*s} \qquad \leftarrow \pi_{t,\tau,j}^d \tag{2l}$$

where:

- S is the $N \times K$ network incidence matrix, whose k^{th} column contains ± 1 for the rows corresponding to the terminal nodes (buses) of circuit k; and zero for the others.
- $f_{t,\tau}$ is the K-dimensional vector of circuit flows $\left\{f_{t,\tau,k}\right\}$.
- $e_{t,\tau}$ is the N-dimensional vector of hydro generation. The energy production $e_{t,\tau,i}$ of each hydro i is in the row of its respective network bus, $n(i)$ (all other values are zero).

- $g_{t,\tau}$ is the N-dimensional vector of thermal generation. The energy production $g_{t,\tau,j}$ of each thermal plant j is in the row of its respective network bus, $n(j)$ (all other values are zero).

- $\hat{d}_{t,\tau}$ and $r^{*s}_{t,\tau}$ are the N-dimensional vectors of load and renewable generation, where each power injection is in the row of its respective network bus (all other values = 0). For the existing renewable plants $(r \in \mathbb{R}_e)$, $r^{*s}_{t,\tau} = \hat{r}^{s}_{t,\tau,r}$. For the candidate renewable plants $(r \in \mathbb{R}_x)$, $r^{*s}_{t,\tau} = \hat{r}^{s}_{t,\tau,r} \times x^{*}_{t,r}$.

Next, we represent Kirchhoff's second law. For the existing circuits $k \in \mathbb{K}_e$, we have:

$$f_{t,\tau,k} = \gamma_k \left(\theta_{t,\tau,F(k)} - \theta_{t,\tau,T(k)}\right) \qquad \forall k \in \mathbb{K}_e \tag{2m}$$

where γ_k is the susceptance of circuit k; $\theta_{t,\tau,F(k)}$ and $\theta_{t,\tau,T(k)}$ are the nodal voltage angles at (respectively) the "from" and "to" terminal buses of circuit k, represented as $F(k)$ and $T(k)$.

For the candidate circuits, the second law is represented as the following constraint:

$$\left|f_{t,\tau,k} - \gamma_k(\theta_{t,\tau,F(k)} - \theta_{t,\tau,T(k)})\right| \leq \Delta^{*}_{t,k} \ \left(= M_k \left[1 - x^{*}_{t,k}\right]\right) \ \leftarrow \ \pi^{\gamma}_{t,\tau,k} \tag{2n}$$

where M_k is a "big M" parameter. We can see that if the candidate circuit is built in the current Benders iteration ($x^{*}_{t,k} = 1$), constraint (2n) becomes equal to equation (2m) of the existing circuits. Conversely, if $x^{*}_{t,k} = 0$, constraint (2n) is relaxed.

Finally, the circuit flow limits are represented as:

$$\left|f_{t,\tau,k}\right| \leq \overline{f}_k \qquad \forall k \in \mathbb{K}_e \tag{2o}$$

$$\left|f_{t,\tau,k}\right| \leq \overline{f}^{*}_{t,k} \ \left(= \overline{f}_k x^{*}_{t,k}\right) \quad \forall k \in \mathbb{K}_x \qquad \leftarrow \ \pi^{\overline{f}}_{t,\tau,k} \tag{2p}$$

Conditioned inflow scenarios for t+1

For simplicity of presentation, we show a multivariate $AR(1)$ model. In practice, SDDP uses a multivariate periodic autoregressive ($PAR(p)$)) model with up to six past time stages:

$$\frac{\left(a^{l}_{t+1,i} - \hat{\mu}_{t+1,i}\right)}{\hat{\sigma}_{t+1,i}} = \hat{\rho}_{t,i} \times \frac{\left(\hat{a}^{s}_{t,i} - \hat{\mu}_{t,i}\right)}{\hat{\sigma}_{t,i}} + \sqrt{1 - \hat{\rho}^{2}_{t,i}} \times \hat{\xi}^{l}_{t,i} \qquad \leftarrow \pi^{a}_{t,i} \tag{2q}$$

where the parameters $\left\{\hat{\mu}_{t,i}, \hat{\sigma}_{t,i}, \hat{\rho}_{t,i}\right\}$ are respectively the mean, standard deviation and serial correlation of the lateral inflow to hydro i in stage t. Spatial dependence is represented through a correlation matrix for the sampled residuals $\left\{\hat{\xi}^{l}_{t,i}\right\}$.

Note: For clarity of presentation, the stochastic streamflow models are shown explicitly. In the actual SDDP implementation, they are represented implicitly.

Future cost functions
As it is well known in SDDP, the FCFs are represented implicitly by a set of hyperplanes.

$$\alpha^{l}_{t+1} \geq \sum_i \hat{\varphi}^{hp}_{t+1,i} \times v_{t+1,i} + \sum_i \hat{\varphi}^{ap}_{t+1,i} \times a^{l}_{t+1,i} + \ \hat{\varphi}^{op}_{t+1} \qquad \forall p = 1, \ \ldots, \ \mathcal{P}; l = 1, \ \ldots, \ L \tag{2r}$$

1.5 Benders cut to the investment module

As it is also well known, the SDDP algorithm is composed of three steps: (i) backward recursion; (ii) forward simulation; and (iii) convergence check. Here, we describe a fourth step used in planning models, which is the calculation of marginal capacity information for a new $(M+1)^{\text{th}}$ Benders cut to the investment module in the final probabilistic forward simulation (after convergence has been achieved):

$$w \geq \sum_t \left(\sum_{i \in \mathbb{I}_x} \overline{\mu}_{t,i}^{M+1} x_{t,i} + \sum_{j \in \mathbb{J}_x} \overline{\mu}_{t,j}^{M+1} x_{t,j} + \sum_{r \in \mathbb{R}_x} \overline{\mu}_{t,r}^{M+1} x_{t,r} + \sum_{k \in \mathbb{K}_x} \overline{\mu}_{t,k}^{M+1} x_{t,k} + \ \overline{\mu}_{0t}^{M+1} \right) \tag{3}$$

The Benders cut coefficients are obtained from the multipliers associated to the following constraints: (i) hydro storage and turbined outflow limits (2d) and (2f); (ii) thermal generation capacity limits (2k); (iii) energy production for the renewables (2l); and (iv) Kirchhoff's second law and flow limits for the transmission components (2m) and (2n).

In the operation problem formulation, we indicated in parenthesis the implicit representation of the investment decision variable (e.g., constraint (2d), reproduced below):

$$v_{t+1,i} \leq \overline{v}_i^* \left(= \overline{v}_i \times x_{t,i}^* \right) \tag{4}$$

This implicit representation is useful for the coefficient calculations, as shown next:

$$\overline{\mu}_{t,i}^{M+1} = \frac{1}{S} \sum_s \left(\overline{v}_i \times \pi_{t,i}^{\overline{v}s} + \overline{u}_i \times \pi_{t,i}^{\overline{u}s} \right) \tag{5a}$$

$$\overline{\mu}_{t,j}^{M+1} = \frac{1}{S} \sum_s \sum_\tau \left(\overline{g}_j \times \pi_{t,\tau,j}^{\overline{g}s} \right) \tag{5b}$$

$$\overline{\mu}_{t,r}^{M+1} = \frac{1}{S} \sum_s \sum_\tau \left(\hat{r}_{t,\tau,r}^s \times \pi_{t,\tau,n(r)}^{\mathrm{ds}} \right) \tag{5c}$$

$$\overline{\mu}_{t,k}^{M+1} = \frac{1}{S} \sum_s \sum_\tau \left(-M_k \times \ \pi_{t,\tau,k}^{\gamma} + \overline{f}_k \times \pi_{t,\tau,k}^{\overline{f}} \right) \tag{5d}$$

$$\overline{\mu}_{0t}^{M+1} = \frac{1}{S} \sum_s \sum_\tau \sum_j c_j g_{t,\tau,j}^{*s} - \sum_{i \in \mathbb{I}_x} \overline{\mu}_{t,i}^{\mathcal{M}+1} x_{t,i}^* - \sum_{j \in \mathbb{J}_x} \overline{\mu}_{gt,j}^{\mathcal{M}+1} x_{t,j}^* - \sum_{r \in \mathbb{R}_x} \overline{\mu}_{t,r}^{\mathcal{M}+1} x_{t,r}^* - \sum_{k \in \mathbb{K}_x} \overline{\mu}_{t,k}^{\mathcal{M}+1} x_{t,k}^* \tag{5e}$$

2 Generation expansion strategy

As mentioned in the Abstract, one of the limitations of an expansion plan is that the construction schedule is defined for the entire study period, and it cannot adapt to changing conditions, for example a severe drought and/or a lower demand growth. For this, we need an expansion *strategy*, similar to the operation strategy calculated by SDDP.

For clarity of presentation, we initially show a simplified expansion strategy formulation, with no construction time and continuous investment decision. We then extend the formulation to represent construction times and, finally, binary investment decisions.

2.1 Formulation 1: no construction time, continuous investment variables

Objective function

In the expansion strategy there are additional state variables for the investment decisions in the previous stage, $\hat{x}_{t-1}$. The objective function is now to minimize the sum of investment costs, immediate cost and expected future costs.

$$\begin{aligned} \alpha_t \left(\hat{v}_t^s, \hat{a}_t^s, \hat{x}_{t-1} \right) = Min \ & \sum_{i \in \mathbb{I}_x} I_i x_{t,i} + \sum_{j \in \mathbb{J}_x} I_j x_{t,j} \sum_{r \in \mathbb{R}_x} I_r x_{t,r} + \sum_{k \in \mathbb{K}_x} I_k x_{t,k} + \sum_j c_j \sum_\tau g_{t,\tau,j} \\ & + \frac{1}{L} \sum_l \alpha_{t+1}^l \end{aligned} \tag{6a}$$

Investment decision constraints

As seen in the investment problem formulation (1), the investment decisions in stage t depend on the previous stage decisions as follows:

$$x_{t,i} \geq \hat{x}_{t-1,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi^x_{t,i} \tag{6b}$$

$$x_{t,j} \geq \hat{x}_{t-1,j} \qquad \forall j \in \mathbb{J}_x \qquad \leftarrow \pi^x_{t,j} \tag{6c}$$

$$x_{t,r} \geq \hat{x}_{t-1,r} \qquad \forall r \in \mathbb{R}_x \qquad \leftarrow \pi^x_{t,r} \tag{6d}$$

$$x_{t,k} \geq \hat{x}_{t-1,k} \qquad \forall k \in \mathbb{K}_x \qquad \leftarrow \pi^x_{t,j} \tag{6e}$$

Water balance equation (same as (2b))

[...]

Storage and turbined outflow limits (no change for existing plants, equations)

Because investments are now decision variables, they are in the LHS of the constraints.

$$v_{t+1,i} - \overline{v}_i \times x_{t,i} \leq 0 \qquad \forall i \in \mathbb{I}_x \tag{6f}$$

$$u_{t,i} - \overline{u}_i \times x_{t,i} \leq 0 \qquad \forall i \in \mathbb{I}_x \tag{6g}$$

Hydropower equations (no change)

[...]

Thermal power constraints (no change for existing plants)

As in the hydro case, investment variables are now on the LHS.

$$g_{t,\tau,j} - \overline{g}_j \times x_{t,j} \leq 0 \qquad \forall j \in \mathbb{J}_x \tag{6h}$$

Renewables

As seen previously, both existing and candidate renewable generation were represented as fixed values in the RHS of the power balance equation. In this case, however, we have to represent explicitly the candidate renewable generation as follows:

$$r^x_{t,\tau,r} - \hat{r}^s_{t,\tau,r} \times x_{t,r} \leq 0 \qquad \forall, r \in \mathbb{R}_x \tag{6i}$$

Bus power balance equation

$$Sf_{t,\tau} + \ e_{t,\tau} + g_{t,\tau} + \ r^x_{t,\tau} = \hat{d}_{t,\tau} - \hat{r}^{\text{es}}_{t,\tau,r} \tag{6j}$$

where $r^x_{t,\tau}$ and $\hat{r}^{\text{es}}_{t,\tau,r}$ are the bus generation vectors for candidate renewables (decision variable) and existing renewables (known values), respectively.

Transmission constraints (no change for existing circuits)

Same investment variable changes as the generation candidates.

$$\left|f_{t,\tau,k} - \gamma_k(\theta_{t,\tau,F(k)} - \theta_{t,\tau,T(k)})\right| - M_k\left(1 - x_{t,k}\right) \leq 0 \qquad \forall k \in \mathbb{K}_x \tag{6k}$$

$$\left|f_{t,\tau,k}\right| - \overline{f}_k x_{t,k} \leq 0 \qquad \forall k \in \mathbb{K}_x \tag{6l}$$

Conditioned inflow scenarios (no change)

[...]

Future cost function

As expected, the FCF now represents the investment state variables.

$$\begin{aligned}\alpha^l_{t+1} \geq &\sum_i \hat{\varphi}^{\text{hp}}_{t+1,i} v_{t+1,i} + \sum_i \hat{\varphi}^{\text{ap}}_{t+1,i} a^l_{t+1,i} + \sum_{i \in \mathbb{I}_x} \hat{\varphi}^{\text{xp}}_{t+1,i} x_{t,i} + \sum_{j \in \mathbb{J}_x} \hat{\varphi}^{\text{xp}}_{t+1,j} x_{t,j} \sum_{r \in \mathbb{R}_x} \hat{\varphi}^{\text{xp}}_{t+1,r} x_{t,r} \\ &+ \sum_{k \in \mathbb{K}_x} \hat{\varphi}^{\text{xp}}_{t+1,j} x_{t,k} + \hat{\varphi}^{\text{op}}_{t+1} \qquad p = 1, \ \ldots, \ P; l = 1, \ \ldots, \ L\end{aligned} \tag{6m}$$

2.1.1 Benders cut for the previous stage

The Benders cut for the previous stage is represented as:

$$\begin{aligned}\alpha_t \geq &\sum_i \hat{\varphi}^{h,P+1}_{t,i} v_{t,i} + \sum_i \hat{\varphi}^{a,P+1}_{t,i} a_{t,i} + \sum_{i \in \mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t,i} x_{t-1,i} + \sum_{j \in \mathbb{J}_x} \hat{\varphi}^{x,P+1}_{t,j} x_{t-1,j} \sum_{r \in \mathbb{R}_x} \hat{\varphi}^{x,P+1}_{t,r} x_{t-1,r} \\ &+ \sum_{k \in \mathbb{K}_x} \hat{\varphi}^{x,P+1}_{t,k} x_{t-1,k} +, \hat{\varphi}^{o,P+1}_t\end{aligned} \tag{7}$$

The cut coefficients are calculated as follows:

$$\hat{\varphi}^{h,P+1}_{t,i} = \pi^h_{t,i} \tag{8a}$$

$$\hat{\varphi}^{a,P+1}_{t,i} = \pi^h_{t,i} + \left(\frac{\hat{\rho}_{t,i}}{\hat{\sigma}_{t,i}}\right) \times \pi^a_{t,i} \tag{8b}$$

$$\hat{\varphi}^{x,P+1}_{t,i} = \pi^x_{t,i} \tag{8c}$$

$$\hat{\varphi}^{x,P+1}_{t,r} = \pi^x_{t,r} \tag{8d}$$

$$\hat{\varphi}^{x,P+1}_{t,k} = \pi^x_{t,k} \tag{8e}$$

$$\begin{aligned}\hat{\varphi}^{o,P+1}_t = &\alpha^*_t - \sum_i \hat{\varphi}^{h,P+1}_{t,i} \hat{v}_{t,i} + \sum_i \hat{\varphi}^{a,P+1}_{t,i} \hat{a}_{t,i} + \sum_{i \in \mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t,i} \hat{x}_{t-1,i} \\ &+ \sum_{j \in \mathbb{J}_x} \hat{\varphi}^{x,P+1}_{t,j} \hat{x}_{t-1,j} \sum_{r \in \mathbb{R}_x} \hat{\varphi}^{x,P+1}_{t,r} \hat{x}_{t-1,r} + \sum_{k \in \mathbb{K}_x} \hat{\varphi}^{x,P+1}_{t,k} \hat{x}_{t-1,k}\end{aligned} \tag{8f}$$

2.2 Formulation 2: construction time, continuous investment variables

For simplicity of presentation, the formulation with construction time shown next only has hydro plants as investment variables. In addition, we assumed that their construction time was three time stages. Given that time stages in the operation problem are just a week or a month long, this is obviously unrealistic: hydro plant construction usually takes five *years*. However, the number of state variables and the logic of their updating in this simplified formulation is very similar to the real models. The reason is that those real models use two types of time stages: years/semesters for investment decisions and weeks/months for operating decisions.

Objective function

The number of state variables for each candidate is equal to their respective construction time (three time stages, in this example):

$$\alpha_t\left(\hat{v}_t^s, \hat{a}_t^s, \hat{x}_{t-1}, \hat{x}_{t-2}, \hat{x}_{t-3}\right) = Min \sum_{i \in \mathbb{I}_x} I_i x_{t,i} + + \sum_j c_j \sum_\tau g_{t,\tau,j} + \frac{1}{L}\sum_l \alpha_{t+1}^l \tag{9a}$$

Investment decision constraints

As seen in the investment problem formulation of section 1, the investment decisions in stage t depend on the previous stage decisions as follows:

$$x_{t,i} \geq \hat{x}_{t-1,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t,i}^x \tag{9b}$$

$$y_{t-1,i} = \hat{x}_{t-1,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t-1,i}^y \tag{9c}$$

$$y_{t-2,i} = \hat{x}_{t-2,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t-2,i}^y \tag{9d}$$

Water balance equation (no change)

[...]

Storage and turbined outflow limits (no change for existing plants)

Because of construction time, the investment decision is known at the beginning of the stage. As a consequence, they go back to the RHS.

$$v_{t+1,i} \leq \overline{v}_i \times \hat{x}_{t-3,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t,i}^{\overline{v}} \tag{9e}$$

$$v_{t,i} \leq \overline{u}_i \times \hat{x}_{t-3,i} \qquad \forall i \in \mathbb{I}_x \qquad \leftarrow \pi_{t,i}^{\overline{u}} \tag{9f}$$

Constraints for hydro generation, thermal, renewables and transmission (no change)

[...]

Future cost function As expected, the FCF now represents the investment state variables for the different time stages.

$$\begin{aligned}\alpha_{t+1}^l \geq & \sum_i \hat{\varphi}_{t+1,i}^{\text{hp}} v_{t+1,i} + \sum_i \hat{\varphi}_{t+1,i}^{\text{ap}} a_{t+1,i}^l + \sum_{i \in \mathbb{I}_x} \hat{\varphi}_{t,i}^{\text{xp}} x_{t,i} + \sum_{i \in \mathbb{I}_x} \hat{\varphi}_{t-1,i}^{\text{xp}} y_{t-1,i} \\ & + \sum_{i \in \mathbb{I}_x} \hat{\varphi}_{t-2,i}^{\text{xp}} y_{t-2,i} + \hat{\varphi}_{t+1}^{\text{op}} \qquad p = 1, \ \ldots, \ P; l = 1, \ \ldots, \ L\end{aligned} \tag{9g}$$

2.2.1 Benders cut for the previous stage

The Benders cut for the previous stage is represented as:

$$\begin{aligned}\alpha_t \geq \sum_i \hat{\varphi}^{h,P+1}_{t,i} v_{t,i} + \sum_i \hat{\varphi}^{a,P+1}_{t,i} a_{t,i} + \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-1,i} x_{t-1,i} \\ + \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-2,i} y_{t-2,i} \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-3,i} y_{t-3,i} + \hat{\varphi}^{o,P+1}_t\end{aligned} \tag{10}$$

The cut coefficients are calculated as follows:

$$\hat{\varphi}^{h,P+1}_{t,i} = \pi^h_{t,i} \tag{11a}$$

$$\hat{\varphi}^{a,P+1}_{t,i} = \pi^h_{t,i} + \left(\frac{\hat{\rho}_{t,i}}{\hat{\sigma}_{t,i}}\right) \times \pi^a_{t,i} \tag{11b}$$

$$\hat{\varphi}^{x,P+1}_{t-1,i} = \pi^x_{t,i} + \pi^y_{t-1,i} \tag{11c}$$

$$\hat{\varphi}^{x,P+1}_{t-2,i} = \pi^y_{t-2,i} \tag{11d}$$

$$\hat{\varphi}^{x,P+1}_{t-3,i} = \overline{v}_i \times \pi^{\overline{v}}_{t,i} + \overline{u}_i \times \pi^{\overline{u}}_{t,i} \tag{11e}$$

$$\begin{aligned}\hat{\varphi}^{o,P+1}_t = \alpha^*_t - \sum_i \hat{\varphi}^{h,P+1}_{t,i} \hat{v}_{t,i} - \sum_i \hat{\varphi}^{a,P+1}_{t,i} \hat{a}_{t,i} - \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-1,i} \hat{x}_{t-1,i} - \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-2,i} \hat{x}_{t-2,i} \\ - \sum_{i\in\mathbb{I}_x} \hat{\varphi}^{x,P+1}_{t-3,i} \hat{x}_{t-3,i}\end{aligned} \tag{11f}$$

2.3 Formulation 3: binary investment variables

Until now we made the simplifying assumption that the investment decisions were continuous variables. This is due to the convexity requirement of the SDDP algorithm which, in theory, would preclude the use of binary variables in the multistage recursion. We have developed two methodologies to address this limitation: (i) hybrid plan/strategy model; and (ii) Stronger convex cuts.

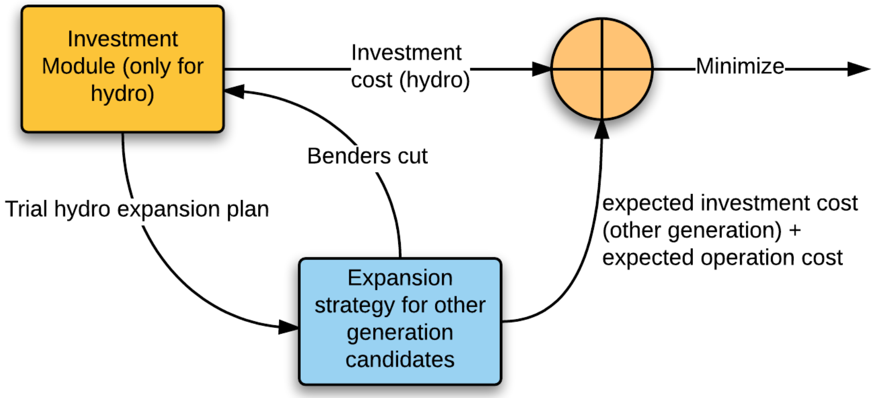
2.3.1 Hybrid Plan/Strategy model

For some types of smaller plants, such as wind, biomass and solar, the use of a continuous approximation is usually reasonable. Even for natural gas plants, depending on system size and on the existing thermal capacity, a continuous relaxation may be acceptable. For large hydro plants, however, it is almost always necessary to use binary variables.

Fortunately, the long construction time of hydro plants, five or more years, provides an "analytical loophole" that allows the representation of their discrete investment decisions within the SDDP framework.

The "loophole" results from the fact that the benefit/cost ratio of building a hydro plant today depends on the probability distribution of future spot prices five years from now (due to the construction period) and later. In the SDDP-based expansion strategy, this future spot price probability distribution is *conditioned* to the values of today's state variables. However, for a construction time of several years, those initial conditions are *not relevant*, that is, the conditioned price probability distribution is equal to the *marginal* distribution. As a consequence, the hydro investment decision (positive or negative) in a given stage will be *the same*, regardless of the system state in that stage. This means that we can represent the hydro binary decisions in the investment module of the expansion *planning* scheme of section 1, while keeping the decisions on the other generation candidates, as part of the expansion *strategy* scheme. The figure below illustrates the hybrid plan/strategy scheme.

Figure 3: Hybrid Plan/Strategy Capacity Expansion Model



2.3.2 Stronger convex cuts

Thomé [4] describes a Lagrangian relaxation methodology for SDDP that produces the strongest convex cut when the one-stage problem is non convex. This has allowed the reduction of duality gaps when binary variables are used.

3 Case Study

3.1 Costa Rica

In this case study, we compare a "traditional" expansion plan and the proposed expansion strategy methodology for the Costa Rica generation system.

Existing generation capacity (2016) comprises: 1,800 MW of hydro; 600 MW of thermal plants; and 180 MW of wind. Annual energy demand is 10,800 GWh.

3.2 Expansion study

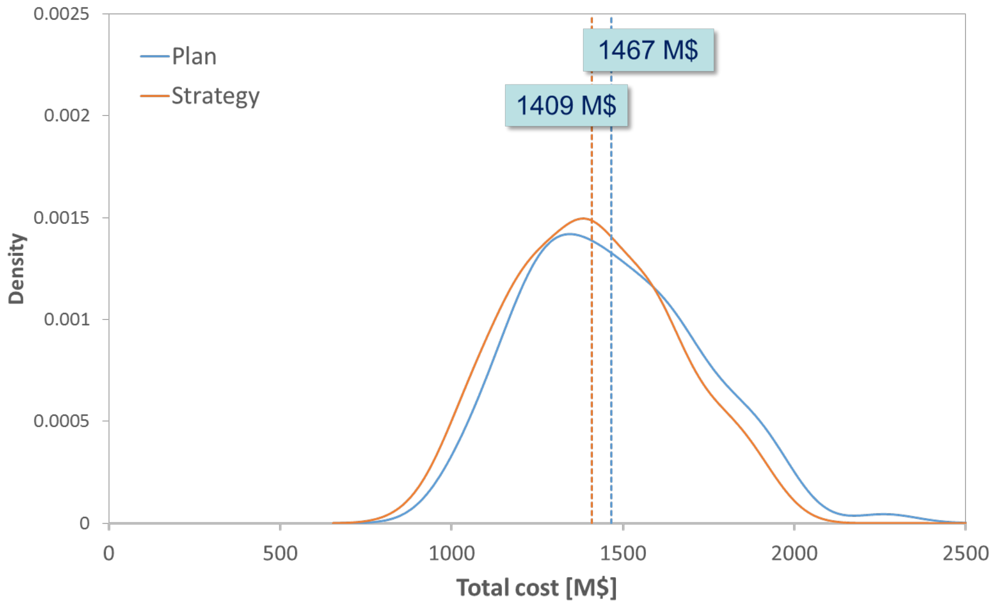
We carried out a ten-year capacity expansion study. The problem parameters are:

- Annual load growth rate: 4%
- Reinforcement candidates: 885 MW of hydro and 1,000 MW of renewable generation
- Annual investment decisions
  - Hydro projects: five-year construction time, binary decision
  - Renewables: one-year construction time, continuous decision
- Monthly operative decisions (SDDP)
  - Monthly load represented by 5 blocks
  - 100 forward scenarios and 30 backward openings in SDDP algorithm

3.3 Study results

Figure (4) shows the probability distribution of total expansion cost (present value of investment + operation costs) using the traditional planning methodology (Plan) and the proposed strategy methodology (Strategy). In the Plan case, investment costs are fixed and operating costs vary with system

Figure 4: Probability distribution of total cost (investment + operation)

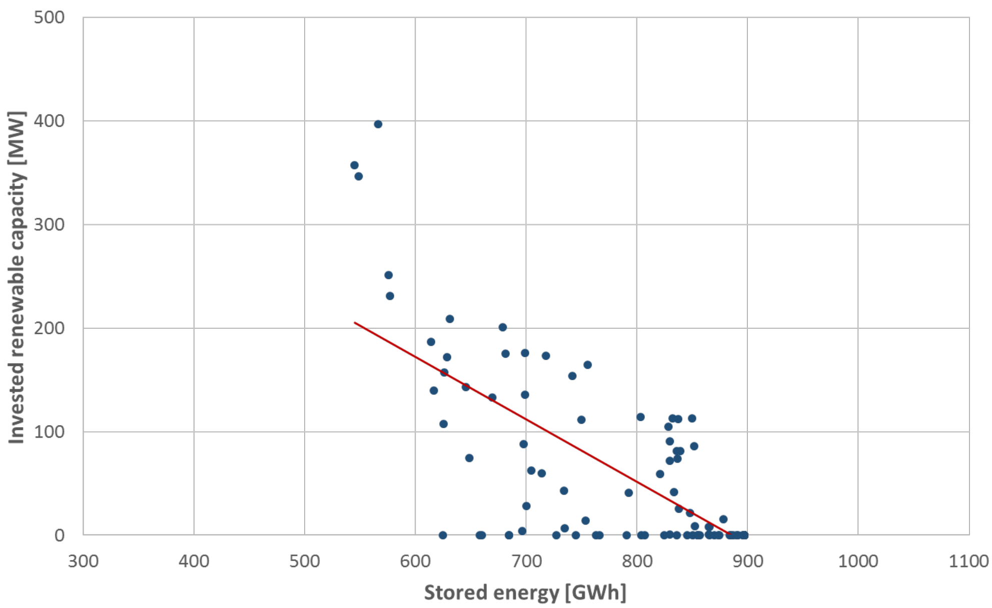


conditions, in particular with the hydrology. In the Strategy case, both investment and operations costs vary with system conditions.

Initially, we observe that the expected total cost of the strategy, 1.409 billion US$, is smaller than the plan's expected cost, 1.467 billion. This result is consistent with the fact that the strategy is more flexible.

Figure (5) shows a scatter plot of investment decisions and system state (reservoir storage) for the year 2020. Again as expected, we see that the strategy invests less in the scenarios with a higher storage level, and vice-versa.

Figure 5: Investment × Stored energy [2020]



3.4 Computational results

Plan

- Number of Benders iterations: 92
- Average number of SDDP iterations for each investment plan: 9

Hybrid Plan/Strategy:

- Number of Benders iterations (planning decisions for hydro): 57
- Average number of SDDPxp (operation + investment for renewables) iterations for each investment plan: 18

4 Conclusions and ongoing research

4.1 Conclusions

The representation of investment decisions in the SDDP algorithm allows the incorporation of uncertainties such as demand growth, fuel costs and hydrology in the decision-making process, which are especially relevant in emerging economies. In addition, it allows the representation of construction times, which is an important advantage of renewable sources, especially in countries with high uncertainty.

4.2 Further research

A recent paper by Ahmed et al [5] showed that the Lagrangian relaxation cuts mentioned in section 2.3.2 are tight when all state variables are binary. This means that we can have an SDDP-based globally optimal strategy with discrete investment variables. We are currently testing this proposed scheme with systems with only thermal and renewable plants. The formulation and algorithm description can be found in [6].

References

[1] B. G. Gorenstin, N. M. Campodónico, J. P. da Costa, and M. V. F. Pereira, "Power system expansion planning under uncertainty," 1993.

[2] M. V. F. Pereira, N. M. Campodónico, B. G. Gorenstin, and J. P. da Costa, "Application of stochastic optimization to power system planning and operation," 1995.

[3] M. V. F. Pereira and L. M. V. G. Pinto, "Multi-stage stochastic optimization applied to energy planning," 1991.

[4] F. S. Thomé, M. V. Pereira, S. Granville, and M. H. C. Fampa, "Non-convexities representation on hydrothermal operation planning using SDDP," 2013.

[5] J. Zou, S. Ahmed, and X. A. Sun, "Nested decomposition of multistage stochastic integer programs with binary state variables," 2016. (Submitted for publication).

[6] F. S. Thomé, C. Metello, C. L. Lara, and S. Granville, "Multistage generation planning model for thermal systems and renewables," 2016.